A New Formulation of the Theorems of Hurwitz, Routh and Sturm

S. BARNETT

School of Mathematics, University of Bradford, Yorkshire, England

[Received 4 March 1971]

It is shown that the Hurwitz determinants associated with a real polynomial of degree n can be obtained from minors of matrices having orders n/2 or (n-1)/2 according as n is even or odd. The method used is based on forming companion matrices of appropriate polynomials, and is extended to calculation of Routh arrays, Sturm sequences and to complex polynomials, thus providing a new formulation of a number of classical theorems.

1. Introduction

FOR A POLYNOMIAL

$$\alpha(\lambda) = \lambda^m + \alpha_1 \lambda^{m-1} + \dots + \alpha_m, \tag{1}$$

with real or complex coefficients, important qualitative problems are to determine the nature and location of the zeros of $\alpha(\lambda)$ without explicit calculation of these zeros. For two or more polynomials a related problem is to find their greatest common divisor (g.c.d.). Such topics have been of interest for a considerable time and, of course, occur in a wide variety of situations arising in a number of branches of applied mathematics. However, it is worth mentioning that the motivation for most of the results to be presented below derived from applications to linear control theory, which in both older and more recent treatments (Rosenbrock, 1970) relies heavily on the theory of polynomials.

One powerful tool for dealing with qualitative problems is provided by Sturm's theorem, which gives the number of distinct real zeros of a real polynomial within a given interval. It is also related to many other well-known results, such as the necessary and sufficient conditions due to Routh and Hurwitz for a real polynomial to be stable (i.e., for all its zeros to have negative real parts). The first general solution of the problem of location of zeros of a complex polynomial in a complex half-plane was in fact derived by Hermite in 1854 using reduction to a sum of squares, and a similar approach by Schur and Cohn evaluated the number of zeros inside the unit circle. Full details and proofs of many location theorems can be found in the book by Marden (1966), and Parks (1963) has given an admirable survey of stability criteria. It is also possible to derive results in this area using resultants and sub-resultants, and relationships between this approach and others are clearly set out by Householder (1968, 1970).

Yet another way of investigating the properties of the polynomial (1) is by using its companion matrix

whose characteristic polynomial is $\alpha(\lambda)$. For example, it is easy to see (as pointed out by MacDuffee (1950)) that if $\beta(\lambda)$ is a second polynomial then det $\beta(C)$ constitutes a resultant of $\alpha(\lambda)$ and $\beta(\lambda)$ —that is, $\beta(C)$ is non-singular if and only if $\alpha(\lambda)$ and $\beta(\lambda)$ are relatively prime. It can be assumed without loss of generality that $\beta(\lambda)$ has degree at most m-1 so that

$$\beta(\lambda) = \beta_0 \lambda^{m-1} + \beta_1 \lambda^{m-2} + \dots + \beta_{m-1}.$$
 (3)

Then not only does $\beta(C)$ have dimensions about half those of Sylvester's resultant, but it is also easy to establish that rows 1, 2, ..., m of

$$\beta(C) = \beta_0 C^{m-1} + \beta_1 C^{m-2} + \dots + \beta_{m-1} I_m, \qquad (4)$$

where I_m is the unit matrix of order *m*, are respectively

$$= [\beta_{m-1}, \beta_{m-2}, ..., \beta_1, \beta_0], xC, xC^2, ..., xC^{m-1}.$$
 (5)

This formulation is interesting because the degree of the g.c.d. of $\alpha(\lambda)$ and $\beta(\lambda)$ (MacDuffee, 1950) and indeed the g.c.d. itself (Barnett, 1970*a*) can be found from $\beta(C)$ in (4) with less effort than is required in the classical case. A straightforward extension holds for the g.c.d. of any number of polynomials (Barnett, 1971*b*), which is directly related to the controllability matrix of a linear multivariable control system, and for which there is no direct classical equivalent. Moreover, the companion matrix approach has produced a simpler reformulation of several other theorems involving or based on resultants, including those on the number of distinct zeros of a polynomial (MacDuffee, 1950), the number of distinct real zeros (Barnett, 1970*c*), Hermite's theorem (Barnett, 1971*a*) and the Schur-Cohn result (Barnett, 1970*b*). It is the purpose of this paper to present further developments of this method, in particular to the calculation of Hurwitz determinants, Routh arrays and Sturm sequences.

In view of the alternatives which have been derived for theorems associated with resultants it is not surprising that a direct relationship between $\beta(C)$ and the Sylvester matrix of $\alpha(\lambda)$ and $\beta(\lambda)$ has been discovered (Barnett, 1971b). It will be convenient for the subsequent development to restate this relationship here, but for the case when $\beta(\lambda)$ in (2) is replaced by a polynomial having the same degree as $\alpha(\lambda)$:

The rows of $\gamma(C)$ are

$$\gamma(\lambda) = \gamma_0 \lambda^m + \gamma_1 \lambda^{m-1} + \ldots + \gamma_m.$$

 $y = [\gamma_m - \gamma_0 \alpha_m, ..., \gamma_2 - \gamma_0 \alpha_2, \gamma_1 - \gamma_0 \alpha_1], \qquad yC, yC^2, ..., yC^{m-1}$ (6) and the Sylvester matrix of $\alpha(\lambda)$ and $\gamma(\lambda)$ is

•	[1	α1	α2	•	•	α_{m-1}	α_m	0	•		•]			
	0	1	α_1	•	•	α_{m-2}	α_{m-1}	α_m	•	•	•			
$S(\alpha, \gamma) =$.	•	•	•	•	•	•	•	•	•	·			
	1.	•		•		•	•	•	•					
	0	0	•	•	•	1	α1	α2	•	•	α_m		(7)	(7)
	0	0	0	•	•	γo	γ1	Y2	•	•	γm			(i)
	•	•	•	•	•	•	•	•	•	•	•			
		•				•		•		•				
	0	Yo	γı	•	•	γ_{m-2}	γ _{m−1}	γm		•				
	γo	γı	Y2	·	•	γ_{m-1}	Υm	0 ·	•	•	•]			
	m	т					. .							
	$\lceil S_1 \rceil$	S_{2}	m				. .							
=	_ <i>S</i> ₃	S_4]m [:]											

It has been shown (Barnett, 1971b) that

$$\begin{bmatrix} I_m & 0\\ J_m K & I_m \end{bmatrix} S(\alpha, \gamma) = \begin{bmatrix} S_1(\alpha, \gamma) & S_2(\alpha, \gamma)\\ 0 & \gamma(C) J_m \end{bmatrix},$$
(8)

where $J_m = [e_m, e_{m-1}, ..., e_1]$, e_i being the *i*th column of I_m , and K is an upper triangular matrix whose precise form is of no importance in what follows. The expression (8) establishes a link between the g.c.d. algorithm referred to earlier (Barnett, 1970a) and a method (Laidacker, 1969) for finding the g.c.d. based on reducing (7) to echelon form. Much more important, however, is the fact that (8) provides the key to the derivation of simpler expressions for the Hurwitz determinants associated with a real polynomial $a(\lambda)$ of degree n. In an earlier paper (Barnett, 1971c) expressions were obtained for the odd Hurwitz determinants when n is even and the even determinants when n is odd, involving minors of a matrix of order n/2 or (n-1)/2 respectively. These were applied to a particular form of the Liénard-Chipart stability criterion rediscovered by Fuller (1957). In Section 2 below the extension to calculation of all the Hurwitz determinants for any real polynomial is given, thus allowing a simplified formulation of the general Liénard-Chipart theorem. Section 3 presents the application of these expressions for the Hurwitz minors to calculation of Routh arrays. It is well known that the Routh stability criterion is equivalent to that of Hurwitz, and it is in fact easy to relate all the elements in a Routh array to (non-principal) minors of the corresponding Hurwitz matrix. By a further application of (8) simplified expressions for these minors are derived, again in terms of matrices of order not more than n/2. It does not seem to be widely known that Routh arrays can be used to obtain Sturm sequences, although the demonstration of this was first published over 70 years ago (Van Vleck, 1899). Fryer (1959) rediscovered this theorem and applied it to find the g.c.d. of two polynomials, and the relationship between this and a companion matrix method (Barnett, 1970a) is explored in Section 4. This is followed by a further application of the results of Section 3 which expresses a Sturm sequence in terms of minors of a single matrix. Finally, in Section 5 the problem of location of zeros of a complex polynomial in a half-plane is considered. It is shown that the theorem of Bilharz (see, e.g. Marden, 1966: 179) reduces to precisely the same form as was found (Barnett, 1971a) when similar methods were applied to Hermite's theorem, thus establishing a new demonstration of the known equivalence of these two results.

2. Hurwitz Determinants

Let

$$a(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \dots + a_n \tag{9}$$

be a polynomial of degree n having real coefficients. Associated with (9) is the Hurwitz matrix of order n which is defined as

$$H = \begin{pmatrix} a_1 & a_3 & a_5 & \cdots & a_{2n-1} \\ 1 & a_2 & a_4 & \cdots & a_{2n-2} \\ 0 & a_1 & a_3 & \cdots & a_{2n-3} \\ 0 & 1 & a_2 & \cdots & a_{2n-4} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & \vdots & a_n \end{pmatrix}$$
(10)

where $a_r = 0$, r > n, and let H_{ii} , i = 1, 2, ..., n denote the leading principal minors of H (the so-called Hurwitz determinants); this notation for leading principal minors will be used generally. The Hurwitz criterion for $a(\lambda)$ to be stable is that all the H_{ii} be positive, and a practical disadvantage is that these minors are not easy to evaluate expect for small values of n. However, the labour involved is reduced by the following:

THEOREM 1. Define the two polynomials

$$f(\lambda) = a_n + a_{n-2}\lambda + a_{n-4}\lambda^2 + \dots$$
(11)

$$g(\lambda) = a_{n-1} + a_{n-3}\lambda + a_{n-5}\lambda^2 + \dots$$
(12)

(i) when n is even, let F be the companion matrix in the form (2) of $f(\lambda)$, which is monic and has degree n/2, and write $L^{(1)} = g(F)J_{n/2}$ and $L^{(2)} = FL^{(1)}$. Then

$$H_{2i-1,2i-1} = L_{ii}^{(1)}, \qquad i = 1, 2, ..., n/2$$
 (13)

$$H_{2i,2i} = (-1)^i L_{ii}^{(2)}, \quad i = 1, 2, ..., (n/2) - 1$$
 (14)

(ii) when n is odd, let G be the companion matrix in the form (2) of $g(\lambda)$, which is monic and has degree (n-1)/2, and write $M^{(1)} = f(G)J_{(n-1)/2}$. Then

$$H_{2i-1,2i-1} = M_{ii}^{(2)}, \qquad i = 1, 2, ..., (n-1)/2$$
 (15)

$$H_{2i,2i} = (-1)^{i} M_{ii}^{(1)} \qquad i = 1, 2, ..., (n-1)/2$$
(16)

where

$$M^{(2)} = \begin{bmatrix} a_1 & a_3 & \dots & a_n \\ & & & 0 \\ M^{(1)} & & 0 \\ & & & \ddots \\ & & & \ddots \\ & & & & 0 \end{bmatrix}.$$
 (17)

Remarks

- (1) Multiplication of a matrix on the right by J of appropriate order simply reverses the order of its columns. If J is removed from the statements (13)-(16), then each *i*th order minor is formed from the last *i* columns of the appropriate matrix and is multiplied by a factor det $J_i = (-1)^{i(i-1)/2}$ (this form was adopted in previous papers, e.g. Barnett (1971*a*,*c*)).
- (2) Construction of g(F) and f (G) is very easy using (5) and (6) respectively. Furthermore in (14), because F is a companion form matrix, the first (n/2)-1 rows of L⁽²⁾ are rows 2, 3, ..., n/2 of L⁽¹⁾.

(3) In both cases
$$H_{nn} = a_n H_{n-1,n-1}$$

Proof. This follows along similar lines to the proof of the main theorem in (Barnett, 1971c), where the expressions (13) and (16) were derived.

(i) when n is even, $f(\lambda)$ in (11) has degree n/2 and $g(\lambda)$ in (12) degree (n/2)-1. To obtain (14), consider $S(f, \lambda g)$, where S is defined in (7). If the rows of H in (10) are numbered 1, 2, ..., n then it is easy to see that $S(f, \lambda g)$ is composed of these rows in the order 2, 4, 6, ..., n, n-1, n-3, ..., 3, 1, so that

$$Z_n H = S(f, \lambda g) \tag{18}$$

where

$$Z_n^T = [e_2, e_4, ..., e_n, e_{n-1}, ..., e_3, e_1].$$
⁽¹⁹⁾

16

The *i*th subresultant of $f(\lambda)$ and $\lambda g(\lambda)$ is defined by

$$R_{i}(f, \lambda g) = \begin{vmatrix} 1 & a_{2} & a_{4} & \dots & a_{4i-2} \\ 0 & 1 & a_{2} & \dots & a_{4i-4} \\ \vdots & & & & \\ 0 & a_{1} & a_{3} & \dots & a_{4i-3} \\ a_{1} & a_{3} & a_{5} & \dots & a_{4i-1} \end{vmatrix}, \qquad i = 1, 2, ..., n/2$$
(20)

where there are *i* rows of each type and as before $a_r = 0$, r > n. Thus $R_i(f, \lambda g)$ is the determinant of the submatrix of $S(f, \lambda g)$ formed by deleting rows and columns numbered 1, 2, 3, ..., (n/2) - i, (n/2) + i + 1, ..., *n*, so in particular $R_{n/2}(f, \lambda g) = \det S(f, \lambda g)$. Inspection of (18) and (20) then readily establishes that $\det Z_{2i}H_{2i,2i} = R_i(f, \lambda g)$ and from (19) it easily follows that $\det Z_{2i} = (-1)^i$. Next, set $\alpha(\lambda) \equiv f(\lambda)$, $\gamma(\lambda) \equiv \lambda g(\lambda)$ and m = n/2 in (8) to obtain

$$\begin{bmatrix} I_{n/2} & 0\\ J_{n/2}K & I_{n/2} \end{bmatrix} S(f, \lambda g) = \begin{bmatrix} S_1(f, \lambda g) & S_2(f, \lambda g)\\ 0 & Fg(F)J_{n/2} \end{bmatrix}.$$
 (21)

Finally, bearing in mind the triangular form of S_1 displayed in (7), application of the Binet-Cauchy theorem to (21) shows that $R_i(f, \lambda g)$ is equal to the *i*th leading principal minor of $Fg(F)J_{n/2}$, which completes the proof.

(ii) when n is odd, the derivation of (15) differs from the foregoing only in minor details. The polynomials (11) and (12) now both have degree (n-1)/2 and the appropriate Sylvester matrix is $S(\lambda g, f)$. A similar argument using slightly modified versions of (7) and (8) shows that $H_{2i-1,2i-1}$ is equal to the *i*th leading principal minor of $f(G_1)J_{(n+1)/2} = M^{(2)}$, where G_1 is the companion matrix of $\lambda g(\lambda)$ so that

$$G_{1} = \begin{bmatrix} 0 & 1 & 0 & . & 0 \\ 0 & & & \\ . & G & & \\ 0 & & & \end{bmatrix} .$$
(22)

Substitution of the coefficients of $f(\lambda)$ into (5) gives as the first two rows of $f(G_1)$

$$r_1 = [a_n, a_{n-2}, ..., a_1], \quad r_2 = r_1 G_1 = [0, a_n - a_1 a_{n-1}, ..., a_3 - a_1 a_2].$$

By applying (6) with $\alpha(\lambda) \equiv g(\lambda)$ and $\gamma(\lambda) \equiv f(\lambda)$ it can then be seen that r_2 is just the first row (z, say) of f(G) preceded by a zero element. The remaining rows $r_2G_1, r_2G_1^2, \dots$ of $f(G_1)$ are therefore, by virtue of (22), zG, zG^2, \dots each preceded by a zero element, and since these are the rows of f(G) this verifies (17).

For (9) to represent a stable polynomial a necessary condition is that all the a_i be positive, and if this is indeed the case then the Liénard-Chipart stability criterion requires calculation of only about half the H_{ii} . Specifically (see Gantmacher, 1959: 221), define the sequences

$$Q_1 = \{a_n, a_{n-2}, a_{n-4}, \ldots\}, \qquad Q_2 = \{a_n, a_{n-1}, a_{n-3}, \ldots\}, Q_3 = \{H_{11}, H_{33}, H_{55}, \ldots\}, \qquad Q_4 = \{H_{22}, H_{44}, H_{66}, \ldots\}.$$

Then a necessary and sufficient condition for $a(\lambda)$ to be stable is that all the members of one of Q_1 , Q_2 and one of Q_3 , Q_4 be positive. Clearly Theorem 1 provides a simplification of this result in that Q_3 and Q_4 can be found using (13), (15), (14), or (16) whichever is appropriate so that the order of the largest determinant to be evaluated is

244

n/2 or (n-1)/2 according as *n* is even or odd. This is the generalization of a result given in a previous paper (Barnett, 1971*c*) in which (13) and (16) were applied to a particular case of the Liénard-Chipart criterion rediscovered by Fuller (1957).

3. Routh Arrays

where

Given two real row vectors

$$c_{0} = [c_{01}, c_{02}, c_{03}, ...]$$

$$c_{1} = [c_{11}, c_{12}, c_{13}, ...]$$
(23)

with $c_{01} \neq 0$, $c_{11} \neq 0$, the Routh array $R\{c_0, c_1\}$ is the name given to the set of rows $c_0, c_1, c_2, c_3, \dots$ defined by

$$c_{i} = [c_{i1}, c_{i2}, c_{i3}, ...], \quad i = 0, 1, 2, ...$$

$$c_{ij} = - \begin{vmatrix} c_{i-2,1} & c_{i-2,j+1} \\ c_{i-1,1} & c_{i-1,j+1} \end{vmatrix} / c_{i-1,1}, \quad i = 2, 3, ...$$
(24)

It will be assumed throughout the rest of this paper that any Routh array under consideration is regular, i.e., $c_{i1} \neq 0$, all *i*. A detailed treatment of non-regular cases can be found in (Gantmacher, 1959: 181). In this section the Routh array associated with $a(\lambda)$ in (9) is defined by taking the vectors in (23) to be the second and first rows of (10) respectively, i.e.

$$\hat{c}_0 = [1, a_2, a_4, \ldots], \qquad \hat{c}_1 = [a_1, a_3, a_5, \ldots],$$
 (25)

and the number of zeros of $a(\lambda)$ in the right half-plane is equal to the number of changes in sign in the first column of $R\{\hat{c}_0, \hat{c}_1\}$. It is convenient to write $R\{\hat{c}_0, \hat{c}_1\}$ (omitting the first row) in the form of a triangular matrix

	[C ₁₁	c_{12}	c_{13}	•	•	•	c_{1n}
	0	c_{21}	c_{22}	•		•	•
T =	0	0	<i>c</i> ₃₁	•	•	•	•
	• *	•	•	•	•	•	•
	0	0	0	•	•	•	c_{n1}

and it is easy to show (Gantmacher, 1959: 191) that

$$\Delta H = T, \tag{26}$$

where H is given by (10) and Λ is lower-triangular with all the elements on its principal diagonal unity. Generalizing the notation used in Section 2 for leading principal minors, let X_{ij} stand for the minor formed from rows 1, 2, ..., i-1, i and columns 1, 2, ..., i-1, j of any matrix X ($j \ge i$). By considering the first i rows in (26) the triangular forms of Λ and T imply

$$H_{ij} = c_{11}c_{21} \dots c_{i-1,1}c_{i,j-i+1}, \qquad j = i, i+1, \dots, n.$$
(27)

In particular taking j = i in (27) gives

$$c_{11} = H_{11}, c_{11}c_{21} = H_{22}, ..., c_{11}c_{21} ... c_{i-1,1}c_{i1} = H_{ii}$$

so that

$$c_{i1} = H_{ii}/H_{i-1,i-1}, \quad i = 2, 3, ..., n.$$
 (28)

Substitution of (28) into (27) gives

$$c_{i,j-i+1} = H_{ij}/H_{i-1,i-1} \qquad i = 2, 3, ..., n,$$

$$j = i, i+1, ..., n.$$
(29)

Equation (28) demonstrates the equivalence of the Hurwitz and Routh stability criteria (the latter being $c_{i1} > 0$, all *i*) and is well-known, but the more general result (29) relating all the coefficients in $R\{\hat{c}_0, \hat{c}_1\}$ to minors of the Hurwitz matrix, whilst equally easy to derive, is not widely recorded. Notice that (28) and (29) show that any row of $R\{\hat{c}_0, \hat{c}_1\}$ can be found individually from minors of a single matrix of order *n* instead of through repeated application of (24), and that regularity of $R\{\hat{c}_0, \hat{c}_1\}$ implies all $H_{ii} \neq 0$.

The application of Routh arrays to computation of Sturm sequences will be developed in the next section, but before this is done a generalization of Theorem 1 to cover all the minors H_{ij} in (29) is presented:

THEOREM 2. For the polynomial $a(\lambda)$ in (9)

(i) when n is even

$$H_{2i-1,j} = L_{i,j-i+1}^{(1)}, \qquad i = 1, 2, ..., n/2, j \ge 2i-1$$
(30)

$$H_{2i,j} = (-1)^{i} L_{i,j-i}^{(2)}, \qquad i = 1, 2, ..., n/2, j \ge 2i$$
(31)

(ii) when n is odd

$$H_{2i-1,j} = M_{i,j-i+1}^{(2)}, \qquad i = 1, 2, ..., (n+1)/2, j \ge 2i-1$$
 (32)

$$H_{2i,i} = (-1)^{i} M_{i,j-i}^{(1)}, \qquad i = 1, 2, ..., (n-1)/2, \ j \ge 2i \tag{33}$$

where $L^{(1)}$, $L^{(2)}$, $M^{(1)}$, $M^{(2)}$ are defined in Theorem 1.

Proof. This proceeds along very similar lines to the proof of Theorem 1. For example, using the same case which was dealt with in detail there, take *n* even and consider again equation (18). It follows that det $Z_{2i}H_{2i,j} = R'_i(f, \lambda g)$ where $R'_i(f, \lambda g)$ is formed from (20) by replacing the last column by elements from the same rows in column (n/2)+j-i of $S(f, \lambda g)$. Again applying the Binet-Cauchy theorem, (21) leads to $R'_i(f, \lambda g) = L^{(2)}_{i,j-i}$ and this establishes (31).

The other cases in Theorem 2 are obtained by similar modifications of the corresponding situations in Theorem 1.

Using equations (30)–(33) with (29) shows that not only the first column but all the elements in $R\{\hat{c}_0, \hat{c}_1\}$ can be obtained in terms of minors of order not more than n/2. The following particular case of this will be useful in the next section.

Corollary. For n even, the elements in
$$R\{\hat{c}_0, \hat{c}_1\}$$
 are given by

$$c_{2i-1,r} = (-1)^{i-1} L_{i,r+i-1}^{(1)} / L_{i-1,i-1}^{(2)} c_{2i,r} = (-1)^{i} L_{i,r+i-1}^{(2)} / L_{ii}^{(1)}$$
(34)

where $L^{(1)}$ and $L^{(2)}$ are defined in Theorem 1.

4. Sturm Sequences

For a real polynomial

$$b(\lambda) = \lambda^N + b_1 \lambda^{N-1} + \dots + b_N \tag{35}$$

a Sturm sequence $\{k_i q_i(\lambda)\}$, i = 0, 1, 2, ... is defined (Lehnigk, 1966: 112) by the remainders, with reversed signs, obtained on successive divisions:

$$\begin{array}{l} q_i(\lambda) = q_{i+1}(\lambda) x_i(\lambda) - q_{i+2}(\lambda), & i = 0, 1, 2, \dots \\ q_0(\lambda) = b(\lambda), & q_1(\lambda) = b'(\lambda) \end{array}$$

$$(36)$$

where $b'(\lambda)$ is the derivative of $b(\lambda)$ with respect to λ , and the k_i are arbitrary positive constants. Sturm's theorem states that the number of distinct real zeros in the interval $\lambda_1 < \lambda < \lambda_2$ is equal to $V(\lambda_1) - V(\lambda_2)$ where $V(\lambda_i)$ is the number of variations in sign in the sequence $q_0(\lambda_i), q_1(\lambda_i), q_2(\lambda_i), \dots$.

Van Vleck (1899) showed that a Sturm sequence can be obtained from alternate rows in the Routh array formed from $b(\lambda)$ and $b'(\lambda)$ by taking as the first two rows $\hat{b}_0 = [1, b_1, b_2, ..., b_N]$ and $\hat{b}_1 = [N, (N-1)b_1, (N-2)b_2, ..., b_{N-1}]$. That is, the polynomials $p_0(\lambda) = b(\lambda)$, $p_1(\lambda) = b'(\lambda)$,

$$p_i(\lambda) = \delta(c_{2i-1,1}\lambda^{N-i} + c_{2i-1,2}\lambda^{N-i-1} + \dots + c_{2i-1,N-i+1}), \quad i = 2, 3, \dots$$
(37)
where

$$\delta = \operatorname{sgn} \left(c_{11} \, c_{21} \, c_{31} \dots \, c_{2i-2,1} \right) \tag{38}$$

and the c_{ij} are the elements in $R\{\hat{b}_0, \hat{b}_1\}$, form a Sturm sequence for $b(\lambda)$. The relationships (37) and (38) were rediscovered by Fryer (1959) who gives details of the procedures to be adopted when a coefficient c_{ii} becomes zero so that the definition in (24) breaks down. In particular, if $b(\lambda)$ has a repeated zero then $b(\lambda)$ and $b'(\lambda)$ have a nontrivial common factor, which leads to a vanishing row in $R\{\hat{b}_0, \hat{b}_1\}$. This is because the process described by (36) is precisely the Euclidean algorithm, apart from the sign reversals which are of no consequence in calculation of the g.c.d. Thus (37) can be used to obtain the g.c.d. of any two polynomials, which without loss of generality may be taken as $\alpha(\lambda)$ and $\beta(\lambda)$ defined in (1) and (3). The coefficients in their g.c.d. $d(\lambda)$ are given by the last non-vanishing row in $R\{\hat{\alpha}, \hat{\beta}\}$, where $\hat{\alpha} = [1, \alpha_1, ..., \alpha_m]$ and $\hat{\beta} = [\beta_0, \beta_1, ..., \beta_{m-1}]$. Specifically, if $d(\lambda)$ has degree k then

$$d(\lambda) = c_{2l,1}\lambda^k + c_{2l,2}\lambda^{k-1} + \dots + c_{2l,k+1}$$
(39)

where l = m - k and $c_{2l,j}$ are the coefficients in row 2l of $R\{\hat{\alpha}, \hat{\beta}\}$. It is of interest to consider the relationship between (39) and the following method for finding the g.c.d. Denote the columns of $\beta(C)J_m$, where $\beta(C)$ is defined in (4), by $y_1, y_2, ..., y_m$. Then it has been shown (MacDuffee, 1950) that $l = \operatorname{rank} \beta(C)J_m$ and (Barnett, 1970a) that $y_1, y_2, ..., y_l$ are linearly independent. Furthermore, if the remaining columns are expressed in terms of this basis:

$$y_i = \sum_{j=1}^{l} x_{ij} y_j, \quad i = l+1, ..., m$$
 (40)

then the monic g.c.d. is

$$\lambda^{k} + x_{l+1,l} \lambda^{k-1} + x_{l+2,l} \lambda^{k-2} + \dots + x_{ml}.$$
 (41)

To relate these two expressions for g.c.d., first apply (34) with f and g replaced by α and β respectively to obtain the coefficients in (39) as

$$c_{2l,p} = (-1)^{l} L_{l,l+p-1}^{(2)} / L_{ll}^{(1)}, \qquad p = 1, 2, ..., k$$
(42)

where

$$L^{(1)} = \beta(C)J_{in}, \qquad L^{(2)} = CL^{(1)}.$$
(43)

S. BARNETT

Next, recalling Remark 2 following Theorem 1, Cramer's rule applied to rows 2, 3, ..., l, l+1 of (40) with i = l+p-1 gives

$$x_{l+p-1,l} = L_{l,l+p-1}^{(2)}/L_{ll}^{(2)}, \quad p = 1, 2, ..., k.$$
(44)

Comparison of (42) and (44) then shows that

$$c_{2l,p} = x_{l+p-1,l}(-1)^{l} L_{ll}^{(2)} / L_{ll}^{(1)}, \qquad p = 1, 2, ..., k,$$
(45)

 $L^{(1)}$ and $L^{(2)}$ being given by (43), and (45) thus demonstrates the equivalence of (39) and (41). The modifications required when $R\{\hat{\alpha}, \hat{\beta}\}$ is not regular are not needed if (40) and (41) are used, and this method is easily extended to deal with more than two polynomials (Barnett, 1971b). Of course in practice the coefficients $x_{l+i,l}$ in (41) would be obtained from (40) by a more direct method than Cramer's rule.

Let us now return to the Sturm sequence (36) associated with the polynomial $b(\lambda)$ in (35). This can be dealt with in a similar fashion to the foregoing, $\alpha(\lambda)$, $\beta(\lambda)$ and m being replaced by $b(\lambda)$, $b'(\lambda)$ and N respectively so that the matrix involved is $L = b'(B)J_N$ where B is the companion matrix of $b(\lambda)$. The coefficients in the Sturm functions (37) are given by (34) as

$$\delta c_{2i-1,r} = \delta (-1)^{i-1} L_{i,r+i-1} / \hat{L}_{i-1,i-1}$$
(46)
where $\hat{L} = BL$, and using (38), (28) and (14),

$$\delta = \operatorname{sgn} (H_{2i-2,2i-2}) = \operatorname{sgn} (-1)^{i-1} \hat{L}_{i-1,i-1}$$

which on substitution into (46) gives

$$\delta c_{2i-1,r} = L_{i,r+i-1} / |\hat{L}_{i-1,i-1}|, \quad r = 1, 2, \dots, N-i+1.$$
(47)

The positive denominators in (47) can be ignored, so that:

THEOREM 3. The polynomials $r_0(\lambda) = b(\lambda)$, $r_1(\lambda) = b'(\lambda)$, $r_i(\lambda) = L_{ii}\lambda^{N-i} + L_{i,i+1}\lambda^{N-i-1} + \dots + L_{iN}$, $i = 2, 3, \dots$

where $L = b'(B)J_N$, form a Sturm sequence for $b(\lambda)$ in (35).

Theorem 3 shows that a Sturm sequence for $b(\lambda)$ can be obtained in terms of minors of a single matrix L, of order N, which is half the order of the matrix used by Van Vleck to derive (37). Notice that b'(B) is easily constructed using a relation of the form (5). It has been shown elsewhere (MacDuffee, 1950) that the rank t of L is equal to the number of real zeros of $b(\lambda)$ and that the number of distinct real zeros is $t-2V_1$, where V_1 is the number of variations in sign in the sequence $1, L_{11}, L_{22}, ..., L_{nt}$ (Barnett, 1970c). In fact the second of these results can be deduced easily from Theorem 3 and Sturm's theorem, since the number of distinct real zeros is $V(-\infty) - V(\infty)$ and $V(-\infty) = t - V(\infty), V(\infty) = V_1$. Similarly, the number of distinct positive real zeros is $V(0) - V(\infty) = V_3 - V_1$ where V_3 is the number of variations in sign in the sequence $b_N, L_{1N}, L_{2N}, ..., L_{1N}$. This corresponds to a recent result of Siljak (1970) expressed in terms of $R\{b_0, b_1\}$. If t < n, multiplicities of real zeros can be found in the usual way (Lehnigk, 1966: 114).

5. Complex Polynomials

Let the polynomial in (9) now have complex coefficients and write

$$a(\lambda) = \lambda^{n} + (a'_{1} + ia''_{1})\lambda^{n-1} + \dots + (a'_{n} + ia''_{n})$$
(48)

248

where the a'_{j} and a''_{j} are real. Following Marden (1966: 179) define the $(2n-1) \times (2n-1)$ matrix

$$P = \begin{bmatrix} a_1' & -a_2'' & -a_3' & a_4'' & a_5' & \dots & (-1)^{n+1}a_{2n-1}' \\ 1 & -a_1'' & -a_2' & a_3'' & a_4'' & \dots & (-1)^{n+1}a_{2n-2}' \\ 0 & a_1' & -a_2'' & -a_3' & a_4'' & \dots & (-1)^{n+1}a_{2n-2}' \\ 0 & 1 & -a_1'' & -a_2' & a_3'' & \dots & (-1)^{n+1}a_{2n-3}' \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \dots & \dots \end{bmatrix}$$
(49)

where $a_r = 0$, r > n. Then provided $a(\lambda)$ has no purely imaginary zeros, the number of zeros of $a(\lambda)$ having positive real parts is equal to the number of variations in sign in the sequence

$$1, P_{11}, P_{33}, \dots, P_{2n-1, 2n-1} \tag{50}$$

provided no member of (50) is zero. Several different but equivalent forms of (49) have been given, and for the case when all the terms in (50) are positive the result is due to Bilharz. It is easy to see that P in (49) has the same form as the Hurwitz matrix (10), the associated Routh array being $R\{\rho_2, \rho_1\}$ where ρ_1 and ρ_2 are the first two rows of (49), so that the sequence (50) can be calculated using (27). However, the method of Section 2 is also immediately applicable by forming polynomials

$$v(\lambda) = \lambda^{n} - a_{1}^{"} \lambda^{n-1} - a_{2}^{'} \lambda^{n-2} + \dots, \qquad w(\lambda) = a_{1}^{'} \lambda^{n-1} - a_{2}^{"} \lambda^{n-2} - \dots$$

from ρ_2 and ρ_1 . Equation (13) then shows that the P_{ii} in (50) are just the leading principal minors of $w(V)J_n$, where V is the companion matrix of $v(\lambda)$. Thus again the minors to be evaluated are reduced in order by about half, and w(V) can be calculated using (5). It is interesting to know that this is precisely the result obtained in an earlier paper (Barnett, 1971a) when the companion matrix method was applied to Hermite's theorem, thus establishing by yet another route the known equivalence of the Hermite and Bilharz criteria (Parks, 1969).

6. Concluding Remarks

It has been shown that the companion matrix approach provides a framework which embraces Hurwitz determinants, Routh arrays and Sturm sequences. This is an extension of previous work and continues the simplified reformulation of a number of classical theorems in the qualitative analysis of polynomials. In particular the orders of any determinants which have to be evaluated can be reduced by about half, so the advantage for hand calculation is obvious. In addition, it is hoped that the new results may lead to improvements over classical algorithms when developed into numerical procedures, and this is currently under investigation. It also seems likely that further relationships involving classical theorems remain to be discovered.

REFERENCES

BARNETT, S. 1970a Linear Algebra and Applications 3, 7–9. BARNETT, S. 1970b Electronics Letters 6, 164–165. BARNETT, S. 1970c I.E.E.E. Trans. autom. Control AC-15, 380–382.

- BARNETT, S. 1971a Linear Algebra and Applications 4, 71-76.
- BARNETT, S. 1971b Proc. Camb. phil. Soc. math. phys. Sci. 70, to appear.
- BARNETT, S. 1971c Proc. Camb. phil. Soc. math. phys. Sci. 70, to appear.
- FRYER, W. D. 1959 I.R.E. Trans. Circuit Theory CT-6, 144-149.
- FULLER, A. T. 1957 Proc. Camb. phil. Soc. math. phys. Sci., 53, 878-896.
- GANTMACHER, F. R. 1959 The theory of matrices, 2. New York: Chelsea.
- HOUSEHOLDER, A. S. 1968 SIAM Rev. 10, 56-66.
- HOUSEHOLDER, A. S. 1970 SIAM Rev. 12, 73-78.
- LAIDACKER, M. A. 1969 Maths Mag. 42, 126-128.
- LEHNIGK, S. H. 1966 Stability theorems for linear motions with an introduction to Liapunov's direct method. Englewood Cliffs; N.J.: Prentice-Hall.
- MACDUFFEE, C. C. 1950 Am. math. Mon. 57, 154-161.
- MARDEN, M. 1966 Geometry of polynomials. Providence, Rhode Island: American Mathematical Society.
- PARKS, P. C. 1963 Proc. Instn mech. Engrs 178 (3M), 7-17.
- PARKS, P. C. 1969 Electronics Letters 5, 55.
- ROSENBROCK, H. H. 1970 State-space and multivariable theory. London: Nelson.
- SILJAK, D. 1970 Proc. I.E.E.E. 58, 1370-1371.
- VAN VLECK, E. B. 1899 Ann. Math. (2nd Ser.) 1, 1-13.